# Finding Mistakes in Basketball Stories: Text 11

Thanks very much for helping with our project! Just to give some background, lots of people are excited by the potential of “deep learning” techniques in artificial intelligence, which have been used in self-driving cars and Google Translate, for example. However deep learning has not been very successful in our research area of “Natural Language Generation”, which is about getting computers to write stories from data.

We are trying to develop better ways of using deep learning in computer systems that write stories from data, and as part of this project we’re trying to figure out how to best detect accuracy problems in computer-generated stories. I.e., cases where the stories say things that aren’t true. We need to find these problems before we can fix them.

As part of this work, we have developed some guidelines for marking up mistakes in computer-generated summaries of basketball games, based on box scores from these games. Focusing on cases where the reports are wrong, not grammar or spelling errors.

What we’d like you to do is to mark-up mistakes in a computer-generated basketball summary, using our guidelines. We’re interested in the mistakes that you find, and also in whether the guidelines make sense. If you find that the guidelines are confusing or don’t work, please let me know!

We’d also like to know some basic information about you

Following this intro, you’ll see

* our guidelines on pages 2 and 3
* an example basketball story which we have marked up, on pages 4 and 5
* the story we would like you to mark up, on page 6. We give links to basketball-reference.com for the box score information as well as the season schedule for each team (such that you can find other games which might be mentioned in the text).
* Space for any additional comments you may have and page 7.

Please mark up the game summary on page 6 in a similar way to the example on pages 4 and send us the document.

Thanks for your help!

Ehud Reiter

## Mark-up guidelines

We’ve given you some basketball game stories produced by “deep learning” AI systems, and also links to box score information on basketball-reference.com about the games (the stories focus on box scores, they don’t talk about individual goals, penalties, etc). We’ve also given links to each teams’ season schedule for season information (some of the stories say where the next game will be). Please read through the stories and mark up cases where

* numbers are wrong
* names (players, teams, cities, etc) are wrong
* words are wrong
* context means people will misunderstand a sentence
* facts which are not checkable
* other cases where the story says something which is not true

We give more information below about these types of errors.

Please mark up the wrong numbers, names, etc by putting them in red. If you’re colour-blind, you can underline them instead. Also please add a note below the story for each mistake; the note should explain the mistake and say which type it is. There is an example on pages 4 and 5.

### Number mistakes

Numbers mistakes are incorrect numbers. For example

* “10-point victory” when margin of victory was 11 pts
* “six players reached double figures” when only four players did so.

Please mark-up the wrong number by putting it in red or underlining it. It doesn’t matter whether the number is digits (such as 10) or written as a word (such as six).

### Name mistakes

Name mistakes are errors in things that have names. This includes people, cities, teams, stadiums, and days of week. If a word (other than “I”) is always capitalised, it is probably a name. For example

* “on Monday” when game was played on Wednesday
* “Talking Stick Resort Arena” when game was played in US Airways Arena.

Please mark-up the wrong name by putting it in red or underlining it.

### Word mistakes

Word mistakes are incorrect or inappropriate words which are not names or numbers. For example

* “out-scored the Suns” when the Suns had a higher score in this period
* “off the bench” for a player who was on the starting team
* “strong first half” when team did poorly in first half

Please mark-up the wrong word by putting it in red or underlining it.

We treat mistakes in fixed phrases such as “off the bench” as word mistakes (the AI systems treat fixed phrases in the same way as they treat words).

The appendix defines some words and phrases which appear in descriptions of basketball games. It is incomplete, let me know if you have any suggestions or corrections for it.

### Context mistakes

Context mistakes occur when people reading a sentence are likely to misinterpret it because of its context, even if the sentence is literally true. For example

* “The Suns had six players reach double figures in points. Mike Conley led the way with 24 points.” This is a context error because Conley played for the other team (not the Suns). Ie, Conley did score 24 points, but the context implies he played for the Suns, which is wrong.

For the mark-up, try to find the thing which will be misinterpreted (as above), and put it in red or underline it.

### Facts which are not checkable

Some facts will not be practical to check. We do not expect you to look forward or back further than 4 games to check a statement. For example:

* Please do check: [not in example game] “It was his second double-double in a row, …”. You can find the players per-game history by clicking on their name in the box score, then selecting the appropriate season. Please also check each teams’ next game details.
* Do not check [not in example game] “The Wizards came into this game as the worst rebounding team in the NBA.”.

For the mark-up, highlight the text which cannot be checked, and put it in red or underline it. Also indicate that it is the “not checkable” category.

### Other mistakes

This includes errors which don’t fit into one of the above categories, and also cases where the story has so many mistakes that it is difficult to individually mark them. For example

* [not in example game] “to give the Suns a commanding Game 1 loss to give the Suns a 118 - 0 record in the Eastern\_Conference 's first playoff series with at least the Eastern\_Conference win in Game 5” when the game is not a playoff game, Suns are not in Eastern Conference, 118-0 is impossible as a team record, etc

Try to mark-up the wrong text by putting it in red or underlining it. We can’t give precise instructions because the “other” category is very broad.

### Notes

If a sentence is repeated, this is not an accuracy mistake, unless of course the sentence contains errors. For example

* “The Sun’s next game will be on Friday, while the Sun’s next game will be on Friday” is a bizarre thing to say, but it is not an accuracy mistake (assuming that the next game will in fact be on Friday).

If there are multiple ways in which you can annotate a sentence for errors, choose the one with the fewest total errors. For Example, when choosing between [not in example game]:

* “Lou Williams led the team in scoring, dropping 30 points, six rebounds and seven assists.”.
* “Lou Williams led the team in scoring, dropping 30 points, six rebounds and seven assists.”.

The first annotation is preferred as the text describes the exact statistics of another player, Solomon Hill.

## Example marked-up story

Game played on 5-Nov-2014, between Memphis Grizzlies and Phoenix Suns.

Main box score data

* <https://www.basketball-reference.com/boxscores/201411050PHO.html>

Other useful data

* Home team season schedule: <https://www.basketball-reference.com/teams/PHO/2015_games.html>
* Visiting team season schedule: <https://www.basketball-reference.com/teams/MEM/2015_games.html>
* Online calendar: <https://www.timeanddate.com/calendar/monthly.html?year=2014&month=11&country=1>

Story, with errors in underlined red

The Memphis Grizzlies ( 5 - 2 ) defeated the Phoenix Suns ( 3 - 2 ) Monday 1 - 2 at the Talking Stick Resort Arena in Phoenix . The Grizzlies had a strong first half where they out - scored the Suns 59 - 42 , to coast to a 10 - point victory in front of their home crowd . The Grizzlies were led by Isaiah Thomas , who scored 15 points ( 4 - 10 FG , 1 - 4 3Pt , 6 - 6 FT ) . He also had six rebounds and five assists in 26 minutes . Eric Bledsoe had 23 points ( 9 - 12 FG , 2 - 2 3Pt , 3 - 4 FT ) , five rebounds and four assists , while Bledsoe added 24 points ( 9 - 14 FG , 2 - 4 3Pt , 3 - 4 FT ) , five rebounds and four assists . The Suns had six players reach double figures in points . Mike Conley led the way with 24 points ( 9 - 14 FG , 3 - 4 3Pt , 3 - 5 FT ) and 11 assists , while Tony Allen chipped in with nine points ( 4 - 6 FG , 1 - 1 FT ) and a pair of assists . The Suns had six players reach double figures in points in this one. Tony Allen had nine points ( 4 - 6 FG , 1 - 1 FT ) and a pair of assists off the bench . The Suns ' next game will be on the road against the Boston Celtics on Friday , while the Suns will be at home against the Portland Trail Blazers on Friday .

LIST OF MISTAKES

5 - 2 - should be 5-0 [number]

Monday – Wednesday [name]

1 - 2 - score was 102-91 [number]

Talking Stick Resort Arena - that is the name now, but in 2014 the Stadium was called US Airways Arena [name]

strong – first half was not strong [word]

outscored – Suns outscored Grizzlies [word]

59-42 – actual score was 46-52 [number]

59-42 – actual score was 46-52 [number]

coast – they had to catch up from behind [word]

10 point – should be 11 point [number]

home – Game was in Phoenix [word]

led – Thomas did not lead the Grizzlies [word]

Isaiah Thomas – Thomas played for Suns, not Grizzlies [context]

six – one rebound [number]

five – two assists [number]

Eric Bledsoe – Bledsoe played for the Suns, not Grizzlies [context]

24, 14, 4 – correct figures for Bledsoe are in previous clause (23, 12, 2) [number]

24, 14, 4 – correct figures for Bledsoe are in previous clause (23, 12, 2) [number]

24, 14, 4 – correct figures for Bledsoe are in previous clause (23, 12, 2) [number]

six – only four Sun players reached double figures [number]

Mike Conley – Conley plays for the Grizzlies, not Suns [context]

Tony Allen – Allen plays for the Grizzlies, not Suns. [context]

Pair – one assist [number]

six – only four Grizzly players reached double figures [number]

off the bench – Allen was a starter [word]

on the road – home game [word]

Boston Celtics – next game is against Sacramento [name]

Portland Trail Blazers – next game is against Sacramento [name]

## Story to be marked-up

Please mark up the following story, links to the statistics for the game on basketball-reference.com, as well as each teams’ season schedule are included below to assist you.

Main box score data

* <https://www.basketball-reference.com/boxscores/201502040ATL.html>

Other useful data

* Home team season schedule: <https://www.basketball-reference.com/teams/ATL/2015_games.html>
* Visiting team season schedule: <https://www.basketball-reference.com/teams/WAS/2015_games.html>
* Online calendar: <https://www.timeanddate.com/calendar/monthly.html?year=2015&month=2&country=1>

STORY

The Atlanta Hawks (41-9) defeated the Washington Wizards (31-19) 105-96 on Tuesday at Philips Arena in Atlanta. The Hawks got off to a quick start in this one, out-scoring the Wizards 29-18 in the first quarter alone. The Hawks were able to out-score the Wizards 28-18 in the third quarter, while the Wizards were able to coast to a victory in front of their home crowd. The Hawks shot 48 percent from the field and 44 percent from the three-point line, while the Wizards finished at 44 percent from the floor and just 25 percent from deep. The Hawks were also able to force the Wizards into 14 turnovers, while committing just 12 of their own. The Hawks were led by the duo of John Wall and Bradley Beal. Wall went 9-for-15 from the field and 3-for-5 from the three-point line to finish with a team-high of 24 points, while also adding nine assists and seven rebounds. It was his second double-double in a row, as he's combined for 54 points and 14 rebounds over his last two games. Notching a double-double of his own, Paul Millsap finished with 11 points (4-7 FG, 2-3 3Pt, 1-2 FT) and 11 rebounds. It was his second double-double in a row, a stretch where he's averaging 21 points and 12 rebounds. Notching a double-double of his own, Bradley Beal finished with 23 points (8-19 FG, 0-4 3Pt, 7-10 FT) and 8 rebounds. He's now averaging 16 points and 6 rebounds on the year. The only other Hawk to reach double figures in points was Al Horford, who finished with 21 points (10-24 FG, 0-1 3Pt, 1-1 FT) and a team-high of 13 rebounds. The Hawks' next game will be at home against the New Orleans Pelicans on Friday, while the Wizards will be at home against the Cleveland Cavaliers on Friday.

LIST OF MISTAKES

<please list mistakes here, as well as marking them up in the story>

* Tuesday 🡪 Wednesday [word]
* alone 🡪 just delete word [word]
  + feels strange to say “in the X quarter/half alone” in an “out-scoring” message, i.e. when the perfomrances of 2 players/teams are being *compared*; “in the X quarter/half alone” seems to be more aplicabale when the performance of a single player/team is being described *in relation to the total performace in the match*:
    - *he recorded 6 blocks in the first quarter alone* ([source](https://books.google.co.uk/books?id=Ks5CDwAAQBAJ&pg=PA2017-IA52&lpg=PA2017-IA52&dq=basketball+%22in+the+first+quarter+alone%22&source=bl&ots=Ym4707YkkI&sig=ACfU3U3oDSEapoMGJgeiFOKzAwNWY35O_Q&hl=en&sa=X&ved=2ahUKEwiJk7rTnfDpAhXJiVwKHcuCAacQ6AEwAHoECAoQAQ#v=onepage&q=basketball%20%22in%20the%20first%20quarter%20alone%22&f=false))
    - *Davis began the game with red eyes, but managed to score 18 points in the first quarter alone.* ([source](https://www.espn.co.uk/nba/recap?gameId=401161368))
    - Roxy Barahman, who had 11 points on 5-of-5 shooting in the first quarter alone. ([source](https://www.thedp.com/article/2020/02/penn-womens-basketball-yale-takeaways-defense-grasela-sterba))
* multiple spaces after full stop (marked with underscores, if this was in the original generate text) 🡪 single (maybe double, style dependent) [orthography]
* Hawks 🡪 Wizards [name]
* Wizards 🡪 Hawks [name]
* while 🡪 but, however [word]
* Wizards 🡪 Hawks [name]
* Hawks 🡪 Wizards [name]
* Wizards 🡪 Hawks [name]
* Hawks 🡪 Wizards [name]
* It was his second double-double in a row 🡪 no double-double in this match, but yes in [the previous match](https://www.basketball-reference.com/boxscores/201502020WAS.html) [context]
* 54 🡪 40 [number] (assumed this as [previous match](https://www.basketball-reference.com/boxscores/201502020WAS.html))
* 14 🡪 8 [number] (assumed this as [previous match](https://www.basketball-reference.com/boxscores/201502020WAS.html))
* Notching a double-double of his own 🡪 no double-double in this match [number]
* 11 rebounds 🡪 only 7 rebounds [number]
* After the sentence “The Hawks were led by the duo of John Wall and Bradley Beal.” the order of details for specific players is John-Wall 🡪 Paul Millsap 🡪 Bradley Beal; The description of details for Paul Millsap should be moved afterwards or at least introduced better, e.g. “on the Hawks’ side…” [cohesion]
* It was his second double-double in a row 🡪 no double-double for Paul Millsap this match [number]
* averaging 21 points 🡪 average 13 points between this and [previous](https://www.basketball-reference.com/boxscores/201502020NOP.html) match [number]
* 12 rebounds 🡪 average 6.5 rebounds between this and [previous](https://www.basketball-reference.com/boxscores/201502020NOP.html) match [number]
* Notching a double-double of his own, Bradley Beal 🡪 second time “Notching a double-double” appears in such a small text; ideally some variation would have occurred [word]
* 8 rebounds 🡪 only 2 rebounds [number]
* The only other Hawk to reach double figures in points was Al Horford 🡪 Jeff teague also had double figures (26) [number]
* The only other Hawk: the preivous player described (Beal) is a Wizards player, so the word “other” is misleading [word]
* 0-1 3Pt 🡪 0-0 3Pt [number]
* New Orleans Pelicans 🡪 [Warriors](https://www.basketball-reference.com/boxscores/201502060ATL.html) [name]
* Cleveland Cavaliers 🡪 [Hornets](https://www.basketball-reference.com/boxscores/201502050CHO.html) [name]
* Friday 🡪 [Thursday](https://www.timeanddate.com/calendar/monthly.html?year=2015&month=2) [word]

## Comments

If you have any comments about this exercise, please write them below.

* I added 2 categories to the annotation: *orthography* and *cohesion*.
* Perhaps easier to use Word’s commenting function than listing errors below text.